

9/19 # 8



OIEP

RAW SEQUENCE LISTING

DATE: 09/26/2002

PATENT APPLICATION: US/10/076,604

TIME: 09:32:53

Input Set : N:\Crf3\RULE60\10076604.raw

Output Set: N:\CRF4\09252002\J076604.raw

SEQUENCE LISTING

1 (1) GENERAL INFORMATION:

2 (i) APPLICANT: White, Tyler R.

3 Damm, Deborah

4 Lesikar, David D.

5 McFadden, Kathleen

6 Garrick, Brett L.

7 (ii) TITLE OF INVENTION: PROTEASE INHIBITOR PEPTIDES

8 (iii) NUMBER OF SEQUENCES: 228

9 (iv) CORRESPONDENCE ADDRESS:

10 (A) ADDRESSEE: Foley & Lardner

11 (B) STREET: 3000 K Street, N.W., Suite 500

12 (C) CITY: Washington

13 (D) STATE: D.C.

14 (E) COUNTRY: USA

15 (F) ZIP: 20007-5109

16 (v) COMPUTER READABLE FORM:

17 (A) MEDIUM TYPE: Floppy disk

18 (B) COMPUTER: IBM PC compatible

19 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

20 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

21 (vi) CURRENT APPLICATION DATA:

C--> 22 (A) APPLICATION NUMBER: US/10/076,604

C--> 23 (B) FILING DATE: 19-Feb-2002

24 (vii) PRIOR APPLICATION DATA:

25 (A) APPLICATION NUMBER: US/09/201,715

26 (B) FILING DATE: 01-Dec-1998

27 (A) APPLICATION NUMBER: US 08/436,555

28 (B) FILING DATE: 08-MAY-1995

29 (viii) ATTORNEY/AGENT INFORMATION:

30 (A) NAME: Pelto, Don J.

31 (B) REGISTRATION NUMBER: 33,754

32 (C) REFERENCE/DOCKET NUMBER: 56324/117

33 (ix) TELECOMMUNICATION INFORMATION:

34 (A) TELEPHONE: (202)672-5300

35 (B) TELEFAX: (202)672-5399

36 (C) TELEX: 904136

37 (2) INFORMATION FOR SEQ ID NO: 1:

38 (i) SEQUENCE CHARACTERISTICS:

39 (A) LENGTH: 57 amino acids

40 (B) TYPE: amino acid

41 (C) STRANDEDNESS: single

42 (D) TOPOLOGY: linear

ENTERED

RAW SEQUENCE LISTING

DATE: 09/26/2002

PATENT APPLICATION: US/10/076,604

TIME: 09:32:53

Input Set : N:\Crif3\RULE60\10076604.raw

Output Set: N:\CRF4\09252002\J076604.raw

```

43      (ii) MOLECULE TYPE: protein
44      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
W--> 45      Xaa Val Cys Ser Glu Gln Ala Glu Xaa Gly Xaa Cys Arg Ala Xaa Xaa
46          1              5              10              15
W--> 47      Xaa Xaa Trp Tyr Phe Asp Val Thr Glu Gly Lys Cys Ala Pro Phe Xaa
48          .              20              25              30
W--> 49      Tyr Gly Gly Cys Xaa Xaa Xaa Xaa Asn Asn Phe Asp Thr Glu Glu Tyr
50          35              40              45
51      Cys Met Ala Val Cys Gly Ser Ala Ile
52          50              55
54 (2) INFORMATION FOR SEQ ID NO: 2:
55      (i) SEQUENCE CHARACTERISTICS:
56          (A) LENGTH: 5 amino acids
57          (B) TYPE: amino acid
58          (C) STRANDEDNESS: single
59          (D) TOPOLOGY: linear
60      (ii) MOLECULE TYPE: protein
61      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
62      Glu Val Val Arg Glu
63          1              5
65 (2) INFORMATION FOR SEQ ID NO: 3:
66      (i) SEQUENCE CHARACTERISTICS:
67          (A) LENGTH: 57 amino acids
68          (B) TYPE: amino acid
69          (C) STRANDEDNESS: single
70          (D) TOPOLOGY: linear
71      (ii) MOLECULE TYPE: protein
72      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
W--> 73      Xaa Val Cys Ser Glu Gln Ala Glu Thr Gly Pro Cys Arg Ala Xaa Xaa
74          1              5              10              15
W--> 75      Xaa Arg Trp Tyr Phe Asp Val Thr Glu Gly Lys Cys Ala Pro Phe Phe
76          20              25              30
W--> 77      Tyr Gly Gly Cys Xaa Gly Asn Arg Asn Asn Phe Asp Thr Glu Glu Tyr
78          35              40              45
79      Cys Met Ala Val Cys Gly Ser Ala Ile
80          50              55
82 (2) INFORMATION FOR SEQ ID NO: 4:
83      (i) SEQUENCE CHARACTERISTICS:
84          (A) LENGTH: 61 amino acids
85          (B) TYPE: amino acid
86          (C) STRANDEDNESS: single
87          (D) TOPOLOGY: linear
88      (ii) MOLECULE TYPE: protein
89      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
90      Glu Val Val Arg Glu Val Cys Ser Glu Gln Ala Glu Thr Gly Pro Cys
91          1              5              10              15
W--> 92      Arg Ala Xaa Xaa Xaa Arg Trp Tyr Phe Asp Val Thr Glu Gly Lys Cys
93          20              25              30
W--> 94      Ala Pro Phe Phe Tyr Gly Gly Cys Xaa Gly Asn Arg Asn Asn Phe Asp

```

RAW SEQUENCE LISTING

DATE: 09/26/2002

PATENT APPLICATION: US/10/076,604

TIME: 09:32:53

Input Set : N:\Crif3\RULE60\10076604.raw

Output Set: N:\CRF4\09252002\J076604.raw

```

95              35              40              45
96      Thr Glu Glu Tyr Cys Met Ala Val Cys Gly Ser Ala Ile
97              50              55              60
99 (2) INFORMATION FOR SEQ ID NO: 5:
100      (i) SEQUENCE CHARACTERISTICS:
101          (A) LENGTH: 57 amino acids
102          (B) TYPE: amino acid
103          (C) STRANDEDNESS: single
104          (D) TOPOLOGY: linear
105      (ii) MOLECULE TYPE: protein
106      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
W--> 107      Xaa Val Cys Ser Glu Gln Ala Glu Xaa Gly Pro Cys Arg Ala Xaa Xaa
108              1              5              10              15
W--> 109      Xaa Xaa Trp Tyr Phe Asp Val Thr Glu Gly Lys Cys Ala Pro Phe Phe
110              20              25              30
W--> 111      Tyr Gly Gly Cys Xaa Gly Asn Arg Asn Asn Phe Asp Thr Glu Glu Tyr
112              35              40              45
113      Cys Met Ala Val Cys Gly Ser Ala Ile
114              50              55
116 (2) INFORMATION FOR SEQ ID NO: 6:
117      (i) SEQUENCE CHARACTERISTICS:
118          (A) LENGTH: 59 amino acids
119          (B) TYPE: amino acid
120          (C) STRANDEDNESS: single
121          (D) TOPOLOGY: linear
122      (ii) MOLECULE TYPE: protein
123      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
124      Val Arg Glu Val Cys Ser Glu Gln Ala Glu Thr Gly Pro Cys Arg Ala
125              1              5              10              15
126      Met Ile Ser Arg Trp Tyr Phe Asp Val Thr Glu Gly Lys Cys Ala Pro
127              20              25              30
128      Phe Phe Tyr Gly Gly Cys Gly Gly Asn Arg Asn Asn Phe Asp Thr Glu
129              35              40              45
130      Glu Tyr Cys Met Ala Val Cys Gly Ser Ala Ile
131              50              55
133 (2) INFORMATION FOR SEQ ID NO: 7:
134      (i) SEQUENCE CHARACTERISTICS:
135          (A) LENGTH: 58 amino acids
136          (B) TYPE: amino acid
137          (C) STRANDEDNESS: single
138          (D) TOPOLOGY: linear
139      (ii) MOLECULE TYPE: protein
140      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
141      Arg Pro Asp Phe Cys Leu Glu Pro Pro Tyr Thr Gly Pro Cys Lys Ala
142              1              5              10              15
143      Arg Ile Ile Arg Tyr Phe Tyr Asn Ala Lys Ala Gly Leu Cys Gln Thr
144              20              25              30
145      Phe Val Tyr Gly Gly Cys Arg Ala Lys Arg Asn Asn Phe Lys Ser Ala
146              35              40              45

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/076,604

DATE: 09/26/2002

TIME: 09:32:53

Input Set : N:\Crif3\RULE60\10076604.raw

Output Set: N:\CRF4\09252002\J076604.raw

```

147      Glu Asp Cys Met Arg Thr Cys Gly Gly Ala
148          50                      55
150 (2) INFORMATION FOR SEQ ID NO: 8:
151   (i) SEQUENCE CHARACTERISTICS:
152       (A) LENGTH: 4 amino acids
153       (B) TYPE: amino acid
154       (C) STRANDEDNESS: single
155       (D) TOPOLOGY: linear
156   (ii) MOLECULE TYPE: protein
157   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
158       Glu Val Val Arg
159           1
161 (2) INFORMATION FOR SEQ ID NO: 9:
162   (i) SEQUENCE CHARACTERISTICS:
163       (A) LENGTH: 79 base pairs
164       (B) TYPE: nucleic acid
165       (C) STRANDEDNESS: single
166       (D) TOPOLOGY: linear
167   (ii) MOLECULE TYPE: DNA (genomic)
168   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
169       TATGAAACAA AGCACTATTG CACTGGCACT CTTACCGTTA CTGTTTACCC CTGTGACAAA      60
170       AGCCGAGGTG TGCTCTGAA                                           79
172 (2) INFORMATION FOR SEQ ID NO: 10:
173   (i) SEQUENCE CHARACTERISTICS:
174       (A) LENGTH: 67 base pairs
175       (B) TYPE: nucleic acid
176       (C) STRANDEDNESS: single
177       (D) TOPOLOGY: linear
178   (ii) MOLECULE TYPE: DNA (genomic)
179   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
180       CTCGGCTTTT GTCACAGGGG TAAACAGTAA CGGTAAGAGT GCCAGTGCAA TAGTGCTTTG      60
181       TTTCATA                                                         67
183 (2) INFORMATION FOR SEQ ID NO: 11:
184   (i) SEQUENCE CHARACTERISTICS:
185       (A) LENGTH: 81 base pairs
186       (B) TYPE: nucleic acid
187       (C) STRANDEDNESS: single
188       (D) TOPOLOGY: linear
189   (ii) MOLECULE TYPE: DNA (genomic)
190   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
191       CAAGCTGAGA CCGGTCCGTG CCGTGCAATG ATCTCCCGCT GGTACTTTGA CGTCACTGAA      60
192       GGTAAGTGCG CTCCATTCTT T                                         81
194 (2) INFORMATION FOR SEQ ID NO: 12:
195   (i) SEQUENCE CHARACTERISTICS:
196       (A) LENGTH: 81 base pairs
197       (B) TYPE: nucleic acid
198       (C) STRANDEDNESS: single
199       (D) TOPOLOGY: linear
200   (ii) MOLECULE TYPE: DNA (genomic)

```

RAW SEQUENCE LISTING

DATE: 09/26/2002

PATENT APPLICATION: US/10/076,604

TIME: 09:32:53

Input Set : N:\Crif3\RULE60\10076604.raw

Output Set: N:\CRF4\09252002\J076604.raw

```

201      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
202      GCACTTACCT TCAGTGACGT CAAAGTACCA GCGGGAGATC ATTGCACGGC ACGGACCGGT      60
203      CTCAGCTTGT TCAGAGCACA C      81
205 (2) INFORMATION FOR SEQ ID NO: 13:
206      (i) SEQUENCE CHARACTERISTICS:
207          (A) LENGTH: 81 base pairs
208          (B) TYPE: nucleic acid
209          (C) STRANDEDNESS: single
210          (D) TOPOLOGY: linear
211      (ii) MOLECULE TYPE: DNA (genomic)
212      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
213      TACGGCGGTT GCGGCGGCAA CCGTAACAAC TTTGACACTG AAGAGTACTG CATGGCAGTG      60
214      TGCGGATCCG CTATTTAAGC T      81
216 (2) INFORMATION FOR SEQ ID NO: 14:
217      (i) SEQUENCE CHARACTERISTICS:
218          (A) LENGTH: 93 base pairs
219          (B) TYPE: nucleic acid
220          (C) STRANDEDNESS: single
221          (D) TOPOLOGY: linear
222      (ii) MOLECULE TYPE: DNA (genomic)
223      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
224      AGCTTAAATA GCGGATCCGC ACACTGCCAT GCAGTACTCT TCAGTGTCAA AGTTGTTACG      60
225      GTTGCCGCCG CAACCGCCGT AAAAGAATGG AGC      93
227 (2) INFORMATION FOR SEQ ID NO: 15:
228      (i) SEQUENCE CHARACTERISTICS:
229          (A) LENGTH: 37 base pairs
230          (B) TYPE: nucleic acid
231          (C) STRANDEDNESS: single
232          (D) TOPOLOGY: linear
233      (ii) MOLECULE TYPE: DNA (genomic)
234      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:
235      CTAGATAAAA GAGAGGTGTG CTCTGAACAA GCTGAGA      37
237 (2) INFORMATION FOR SEQ ID NO: 16:
238      (i) SEQUENCE CHARACTERISTICS:
239          (A) LENGTH: 37 base pairs
240          (B) TYPE: nucleic acid
241          (C) STRANDEDNESS: single
242          (D) TOPOLOGY: linear
243      (ii) MOLECULE TYPE: DNA (genomic)
244      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:
245      CCGGTCTCAG CTTGTTCAGA GCACACCTCT CTTTTAT      37
247 (2) INFORMATION FOR SEQ ID NO: 17:
248      (i) SEQUENCE CHARACTERISTICS:
249          (A) LENGTH: 49 base pairs
250          (B) TYPE: nucleic acid
251          (C) STRANDEDNESS: single
252          (D) TOPOLOGY: linear
253      (ii) MOLECULE TYPE: DNA (genomic)
254      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/076,604

DATE: 09/26/2002
TIME: 09:32:54

Input Set : N:\Crf3\RULE60\10076604.raw
Output Set: N:\CRF4\09252002\J076604.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos.1,9,11,15,16,17,18,32,37,38,39,40
Seq#:3; Xaa Pos.1,15,16,17,37
Seq#:4; Xaa Pos.19,20,21,41
Seq#:5; Xaa Pos.1,9,15,16,17,18,37
Seq#:63; N Pos. 24
Seq#:66; N Pos. 22,26
Seq#:67; N Pos. 22
Seq#:102; N Pos. 103,104,106,107,109,110,112,113
Seq#:102; Xaa Pos.268,269,270,271
Seq#:103; Xaa Pos.35,36,37,38

VERIFICATION SUMMARY

DATE: 09/26/2002

PATENT APPLICATION: US/10/076,604

TIME: 09:32:54

Input Set : N:\Crif3\RULE60\10076604.raw

Output Set: N:\CRF4\09252002\J076604.raw

L:22 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:23 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:45 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:0
L:47 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:16
L:49 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:32
L:73 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:0
L:75 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:16
L:77 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:32
L:92 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:16
L:94 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:32
L:107 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0
L:109 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:16
L:111 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:32
L:885 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:76
L:888 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:76
L:891 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:76
L:894 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:76
L:925 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:78
L:928 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:78
L:931 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:78
L:934 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:78
L:969 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:80
L:972 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:80
L:975 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:80
L:978 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:80
L:981 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:80
L:984 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:80
L:987 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:80
L:990 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:80
L:993 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:80
L:996 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:80
L:1039 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:82
L:1042 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:82
L:1045 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:82
L:1048 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:82
L:1051 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:82
L:1054 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:82
L:1057 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:82
L:1060 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:82
L:1063 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:82
L:1108 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:84
L:1111 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:84
L:1114 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:84
L:1117 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:84
L:1120 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:84
L:1123 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:84
L:1126 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:84
L:1129 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:84

VERIFICATION SUMMARY

DATE: 09/26/2002

PATENT APPLICATION: US/10/076,604

TIME: 09:32:54

Input Set : N:\Crif3\RULE60\10076604.raw

Output Set: N:\CRF4\09252002\J076604.raw

L:1132 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:84
L:1177 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:86
L:1180 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:86
L:1183 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:86
L:1186 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:86
L:1189 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:86
L:1192 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:86
L:1195 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:86
L:1198 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:86
L:1201 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:86
L:1246 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:88
L:1249 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:88
L:1252 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:88
L:1255 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:88
L:1258 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:88
L:1760 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:102 after pos.:144
L:1810 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:103 after pos.:32